



## OPEN ACCESS

## EDITED BY

Helen Higham,  
University of Oxford, United Kingdom

## REVIEWED BY

Charles Vincent,  
University of Oxford, United Kingdom  
Sherry Espin,  
Toronto Metropolitan University, Canada

## \*CORRESPONDENCE

Emily Grossniklaus  
✉ emily77@uw.edu

<sup>†</sup>These authors share first authorship

RECEIVED 15 February 2025

ACCEPTED 06 June 2025

PUBLISHED 20 June 2025

## CITATION

Grossniklaus E, D'Addario A, King A,  
Gallagher TH, Mazor K and White AA (2025)  
Error disclosure: what residents say and what  
patients find effective.  
Front. Health Serv. 5:1577092.  
doi: 10.3389/frhs.2025.1577092

## COPYRIGHT

© 2025 Grossniklaus, D'Addario, King,  
Gallagher, Mazor and White. This is an open-  
access article distributed under the terms of  
the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)  
(CC BY). The use, distribution or reproduction  
in other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Error disclosure: what residents say and what patients find effective

Emily Grossniklaus<sup>1\*†</sup>, Angelo D'Addario<sup>2†</sup>, Ann King<sup>2</sup>,  
Thomas H. Gallagher<sup>3</sup>, Kathleen Mazor<sup>4</sup> and Andrew A. White<sup>5</sup>

<sup>1</sup>Department of Medicine, VA Puget Sound Healthcare System and University of Washington School of Medicine, Seattle, WA, United States, <sup>2</sup>Office of Research Strategy, National Board of Medical Examiners, Philadelphia, PA, United States, <sup>3</sup>Department of Medicine and Department of Bioethics, University of Washington School of Medicine, Seattle, WA, United States, <sup>4</sup>Department of Medicine, UMass Chan Medical School, Worcester, MA, United States, <sup>5</sup>Department of Medicine, University of Washington School of Medicine, Seattle, WA, United States

**Background:** Medical error disclosure to patients is a critical skill that is often not taught effectively in medical training. The Video-based Communication Assessment (VCA) software enables trainees to receive feedback on their error disclosure communication skills. The VCA method also allows examination of the specific types of error disclosure responses that patients value most.

**Objective:** The primary aim of this study was to describe the language medical residents use to disclose a hypothetical harmful medical error, and to determine the language associated with higher ratings by crowdsourced laypeople. A secondary aim of this study was to examine the alignment between error disclosure content recommended by experts and the communication behaviors that contribute to higher layperson ratings of disclosure.

**Methods:** 102 resident physician responses to a case depicting a delayed diagnosis of breast cancer and their crowdsourced ratings were analyzed using thematic content analysis. We assessed the presence of specific themes in response to three sequential video prompts within a clinical case. Linear regressions were then performed for each prompt's response to examine the extent to which each theme predicted overall communication scores from layperson raters.

**Results:** Nearly all ( $N = 92$ , 90.2%) residents provided responses which included either a general apology or a specific apology in at least one of the three prompt's responses, and nearly all ( $N = 98$ , 96.1%) residents provided at least one response expressing a component of empathy. However, only 57.8% of residents openly acknowledged that the care was delayed, and 67.8% expressed a plan to prevent future errors. A few residents used rationalization (5.9%) or minimization (4.9%) behaviors; responses with these behaviors were associated with negative beta-coefficients, although this finding did not reach statistical significance. In a linear regression analysis, the strongest positive associations between resident responses and patient ratings were clustered around expressions of accountability (0.48), personal regret (0.47), apology (0.34), and intentions to prevent future mistakes (0.34).

**Conclusion:** Resident physicians vary in which communication elements and themes they include during error disclosure, missing opportunities to meet patient expectations. While infrequent, some residents employed minimization or rationalization in their responses. Utilizing an assessment and feedback system that encourages responders to include themes layperson raters value most and to omit harmful expressions could be an important feature for future software for error disclosure communication training.

#### KEYWORDS

error disclosure, communication training, graduate medical education, communication skills assessment, crowdsourcing

## Introduction

It is widely acknowledged that physicians should promptly disclose harmful errors to patients, explain how the error occurred, apologize, and provide patients with emotional support (1, 2). Doing so with transparency and caring fulfills ethical and regulatory obligations, minimizes patient suffering, and promotes organizational learning (3). Unfortunately, conversations after errors frequently fall short of patient expectations because they lack the information or empathy sought by patients (4, 5). Physicians report inadequate training as a key contributor to their difficulty with effective error disclosure (6–8). In response, mandates have emerged in the US for resident physicians to receive training on the disclosure of patient safety events (9), and for hospitals to support practicing physicians preparing to disclose an error (10).

To meet these requirements, educators and coaches need to instruct physicians on what to say, how to say it, and what not to say. Current guidance about how clinicians should disclose errors is informed by patient surveys, expert opinion, and limited empiric evidence that open communication and an apology can reduce patients' long-term distress (11–14). However, studies describing physicians' actual disclosure language and its relationship to disclosure efficacy are lacking (15). This gap limits the design of coaching and teaching tools that aim to narrow the difference between doctors' tendencies and the communication behaviors desired by patients.

The legal and emotional sensitivity of real-life error disclosure makes it difficult to observe or record these conversations for research. Instead, efforts to describe physician disclosure language have used surveys in which physicians are asked to select which multiple choice option is closest to how they would respond to a hypothetical error (4, 16), analysis of free text written responses to hypothetical errors (17), focus groups with care providers about their approach to communicating diagnostic delays (18), and assessments of physician performance in error disclosure simulations with standardized patients (19–22). While simulations maybe realistic for eliciting communication skills, research to date has not examined the relationship between observed communication behaviors and patients' assessments. We sought to describe this relationship through novel analysis of audio recordings of resident physician responses to a simulated medical error portrayed in the Video-based Communication Assessment (VCA). The primary purpose of this study was to

describe the language residents would use to disclose an error, and to determine the association of those language choices with layperson ratings of their disclosure skills.

The VCA is a smartphone-based software tool for physicians to practice their error disclosure skills and receive feedback on their responses (23). The VCA app describes a medical case, presents a video prompt of a simulated patient, and directs the user to audio-record what they would say in response to the patient prompt in the scenario. Responses are reviewed by panels of 8–10 crowdsourced laypeople, who provide ratings on key domains of error disclosure quality on a five-point scale and qualitative responses on what would constitute the ideal response. Users receive feedback reports with a numerical summary rating and learning points derived from raters' comments (24, 25). These reports are designed so that users can self-direct learning of communication skills, and provides them with high-quality, actionable feedback, without requiring the high level of resources needed for standardized patient interactions (26). This aligns with the experiential learning framework, as it enables trainees to build their knowledge through experience and feedback (27).

In a recent multi-center randomized trial, resident physicians practiced error disclosure with the VCA and were randomized to either receive crowdsourced feedback on their performance or not. The study found that residents who received feedback had significantly higher error disclosure communication skills ratings from laypeople during their subsequent engagement with the VCA (26). Prior studies of the VCA's validity for rating physician skill (25, 26), feasibility (28), and providing actionable advice have been encouraging (23, 24). Laypeople have been shown to provide reliable and consistent VCA ratings when compared to panels of patients who have experienced harm from medical care (29). However, it is unknown whether layperson ratings of residents' communication skills align with behaviors recommended by physician faculty experts, who have traditionally assessed trainee skills in simulated patient encounters. Thus, a secondary purpose of this study was to evaluate alignment between layperson ratings and the behaviors recommended by experts.

## Materials and methods

Using an existing collection of resident audio recordings collected with the VCA, we (1) conducted a thematic analysis to

identify the communication behaviors present in the responses, and (2) performed multiple linear regression to determine which of these behaviors predicted layperson quantitative ratings. Data were derived from resident physician responses to a case depicting a delayed diagnosis of breast cancer used in two studies evaluating the efficacy of the VCA for error disclosure skill development. These studies included a 2022 pre/post study (25) and a 2024 randomized trial (26) with a combined cohort of 102 Internal medicine and family medicine residents who completed the case during our period of analysis. The University of Washington IRB approved all study procedures.

## Participants, VCA case

We recruited resident physicians from 6 Internal Medicine and Family Medicine residency programs across 5 US states to participate in classroom trials of the VCA; details of the VCA app and methods of these studies have been previously published (25, 26); essential components are summarized here. The participants were given protected time during a didactic session on error disclosure to respond to two VCA cases, including the one analyzed here. The case depicts an error in which a mammogram report concerning for breast cancer was missed by the care team for a year, leading to a delayed diagnosis. The written scenario description and the script of the video prompt are shown in the grey bubbles of Figures 1–3. The case includes three sequential segments (which will be termed vignettes) in which the actor depicts a 58-year-old patient (a) asking whether

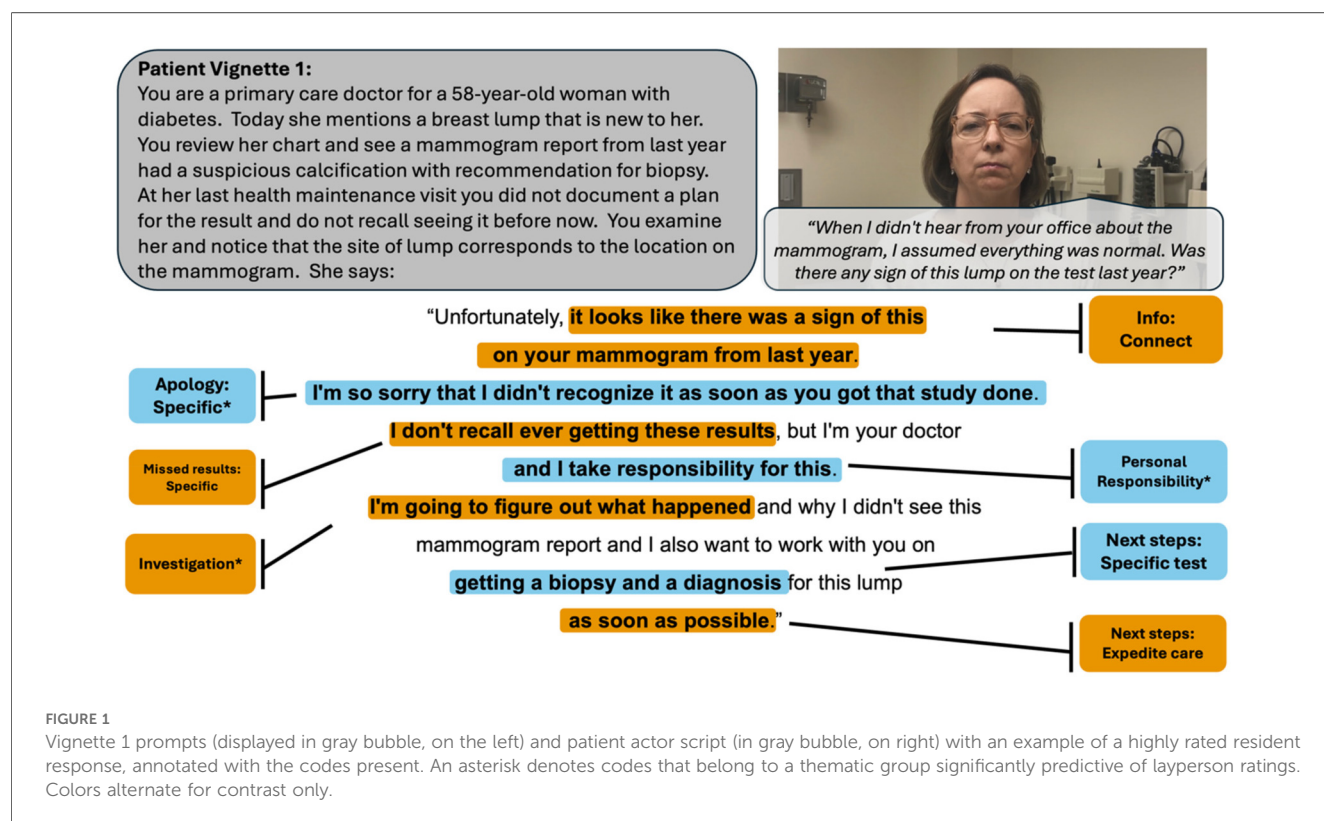
a breast lump was visible on the mammogram a year ago, (b) expressing anxious and tearful reactions to the revelation that the diagnosis was delayed by a medical error, and (c) expressing mistrust in the medical system. Residents entered the app through a personal login and password and provided audio responses to each vignette through the VCA software.

## Crowdsourced ratings

As previously described, panels of 8–10 US-based, English-speaking adults were recruited on Amazon's Mechanical Turk (MTurk) to rate physicians' recorded audio responses (25, 26). Laypeople were chosen as raters because they can be recruited rapidly and affordably, and represent the broad patient population potentially exposed to medical error. Laypeople rated each response on 6 items covering domains related to accountability, honesty, apology, empathy, caring, and overall response. Items used a 5-point scale anchored with the labels poor, fair, good, very good, and excellent. We sought at least 6 raters per response after removing raters with incomplete responses or evidence of inattention (30).

## Qualitative data analysis

We analyzed transcripts of each physician's response using a thematic approach. To develop a codebook (Supplementary Appendix 1), one co-author (EJG) first reviewed 15 responses to guide initial inductive creation of codes. Then, codes were



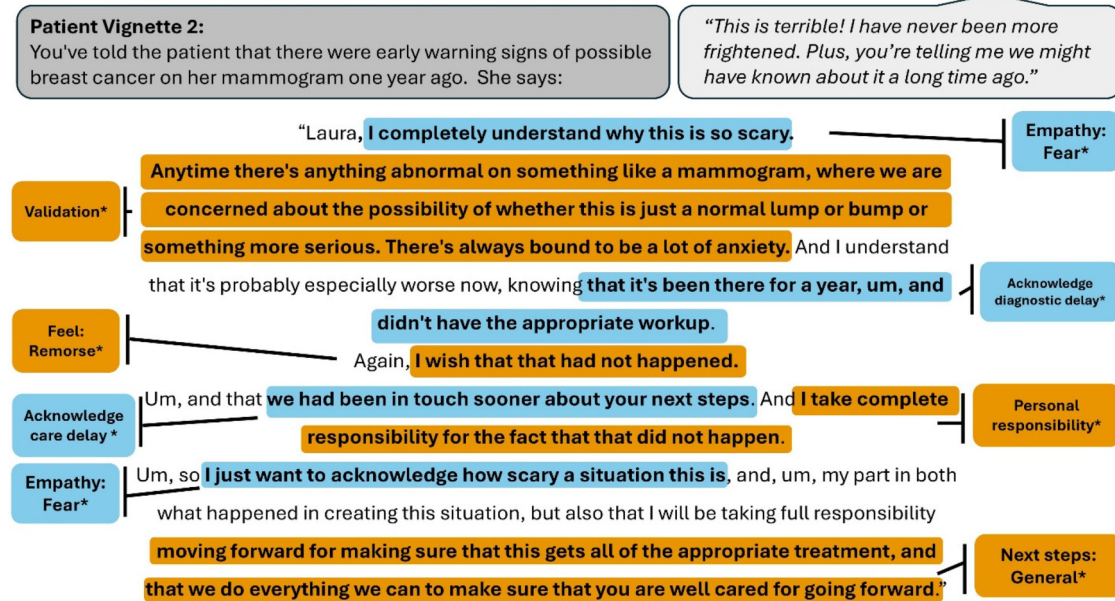


FIGURE 2

Vignette 2 prompts (displayed in gray bubble, on the left) and patient actor script (in gray bubble, on right) with an example of a highly rated resident response, annotated with the codes present. An asterisk denotes codes that belong to a thematic group significantly predictive of layperson ratings. Colors alternate for contrast only.

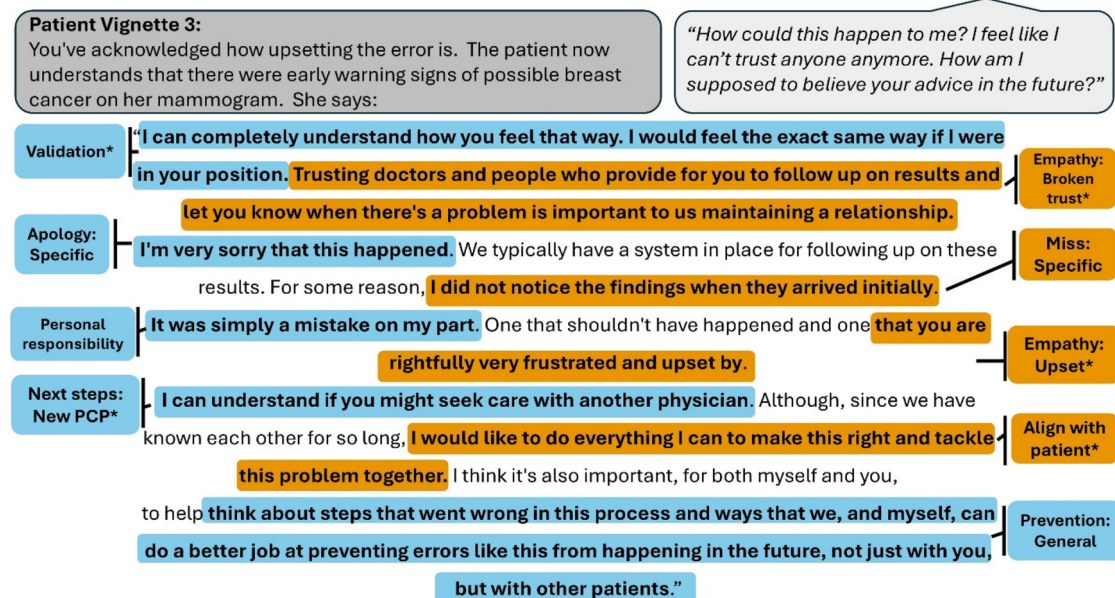


FIGURE 3

Vignette 3 prompts (displayed in gray bubble, on the left) and patient actor script (in gray bubble, on right) with an example of a highly rated resident response, annotated with the codes present. An asterisk denotes codes that belong to a thematic group significantly predictive of layperson ratings. Colors alternate for contrast only.

reviewed by a set of co-authors with expertise in error disclosure communication (THG, KM, AAW) and the CANDOR error disclosure process (31). This deductive approach drew attention to the nature of the error (diagnostic delay caused by an unclear

combination of physician and system factors), and the patient's primary emotional reactions manifested by the actor (fear, mistrust). We recharacterized 2 codes and added 7 codes, yielding a total of 38 codes in 15 thematic groups. Of the thematic



groups, the panel recommended that error disclosure communication avoid 3 (minimization, rationalization, jargon), possibly include 2 (using patient name and personal feelings) and address the remaining 10 (acknowledge delays, acknowledge missed results, apology, accept responsibility, empathy, alignment, next steps, system improvements, financial, address mammogram results). Two co-authors (ADA and EG) reviewed and coded each transcript independently using Atlas.ti v24 (Atlas.ti Scientific Software Development GmbH; Berlin, Germany). Coding was reviewed to achieve consensus; areas of disagreement were resolved by group discussion with a third co-author (AAW).

Quantitative data analysis

We grouped codes with similar themes for each vignette (e.g., “*acknowledge diagnostic delay*” and “*acknowledge delay of care*” were grouped together under the thematic group “*Acknowledge delays*”). We then created binary variables indicating whether a physician’s response to a single vignette was coded for any code under a given thematic group. Responses with one or multiple codes within a theme were marked as “1”, whereas responses with none were marked as “0”. Binary coding was chosen to represent the presence or absence of each thematic group in a response, rather than emphasizing individual elements or their frequency.

We computed a mean overall rating score for each physician’s response by aggregating lay rater’s ratings across items and raters for each of the three vignettes. We chose to present aggregate scores of layperson ratings, as this scale has demonstrated high internal reliability in our prior operational research, and learners typically receive feedback in this aggregated matter. This aggregation yielded three separate communication scores for each physician (one per vignette). We then performed three separate multiple linear regressions to examine the extent to which each thematic code group predicted communication scores from lay raters for each vignette. Though the three vignettes were structured as an unfolding sequence within a broader case, we analyzed them separately to preserve time-specific contextual variation in communication behaviors that could be lost through aggregation. Certain communication behaviors, like offering an apology or acknowledging uncertainty, may have varying levels of salience depending on the phase of the encounter. Analyzing each vignette individually allowed us to capture these patterns and prevent losing granular insight that might be obscured by aggregation. Each regression model included the binary variables representing the thematic groups as predictors and the aggregated communication scores as the dependent variable. Statistical analyses were conducted using R version 4.1.

Results

Participant characteristics

Characteristics of the resident physician participants and layperson raters are presented in Table 1. The mean overall

TABLE 1 Resident participant and layperson rater characteristics.

Characteristic	N	%
(a) Resident physician characteristics (N = 102)		
Gender		
Female	45	44.1%
Male	52	51.0%
Non-binary	1	1.0%
Prefer not to say	4	3.9%
Specialty		
Internal Medicine	89	87.3%
Family Medicine	13	12.7%
(b) Layperson rater characteristics (N = 234)		
Gender		
Female	90	38.5%
Male	143	61.1%
Prefer not to say	1	0.4%
Age (years)		
18–24	10	4.3%
25–34	86	36.8%
35–44	66	28.2%
45–54	37	15.8%
55–64	28	12.0%
65+	7	3.0%
Education		
Some high school	1	0.4%
Graduated high school or equivalent	24	10.3%
Some college, no degree	29	12.4%
Associate degree	14	6.0%
Bachelor’s degree	129	55.1%
Graduate degree	37	15.8%

layperson scores for responses were 3.25 (SD = 0.49, IQR = 0.75) for vignette 1, 3.37 (SD = 0.43, IQR = 0.51) for vignette 2 and 3.51 (SD = 0.39, IQR = 0.48) for vignette 3.

Communication elements present in resident responses

The prevalence of communication themes across the entire case and representative quotes are presented in the text below. Themes varied significantly across vignettes; counts and percentages per vignette are shown in Table 2.

Explanatory clinical information

Virtually all residents (N = 101, 99.0%) provided some explanation of clinical information in at least one of their responses. Most (N = 80, 78.4%) residents referenced the findings from the prior mammogram in at least one of the three vignettes. Many (N = 64, 62.8%) also connected prior findings to the current concerns. For example, one resident said, “*On your mammogram last year, there was actually an indication of a suspicious calcification*” and another said “*And, that is the same area that you are now feeling this lump.*”

TABLE 2 Results table showing the number and percentage of resident responses that included each code for every vignette prompt.

Group Code	Vignette 1		Vignette 2		Vignette 3	
	N (%) Present	$\beta$ (SE)	N (%) Present	$\beta$ (SE)	N (%) Present	$\beta$ (SE)
<b>Acknowledge Delays</b>						
Acknowledge diagnostic delay	16 (15.7)	–	35 (34.3)	–	2 (2.0)	–
Acknowledge care delay	10 (9.8)	–	14 (13.7)	–	1 (1.0)	–
Grouping Total	26 (25.5)	0.17 (0.09) <sup>+</sup>	46 (45.1)	0.19 (0.08)*	3 (2.9)	–0.16 (0.22)
<b>Acknowledge missed results</b>						
Missed results: General	26 (25.5)	–	16 (15.7)	–	5 (4.9)	–
Missed results: Specific	45 (44.1)	–	10 (9.8)	–	2 (2.0)	–
Grouping Total	63 (61.8)	0.03 (0.08)	26 (25.5)	–0.09 (0.09)	7 (6.9)	0.03 (0.15)
<b>Apology</b>						
Apology: General	16 (15.7)	–	25 (24.5)	–	13 (12.8)	–
Apology: Specific	50 (49.0)	–	54 (52.9)	–	29 (28.4)	–
Grouping Total	60 (58.8)	0.34 (0.09) <sup>***</sup>	71 (69.6)	0.11 (0.09)	38 (37.3)	0.03 (0.08)
<b>Accept responsibility</b>						
Apology with accountability	8 (7.8)	–	8 (7.8)	–	9 (8.8)	–
Personal responsibility	25 (24.5)	–	20 (19.6)	–	13 (12.8)	–
Disclose error	8 (7.8)	–	15 (14.7)	–	37 (36.3)	–
Grouping Total	30 (29.4)	0.48 (0.09) <sup>***</sup>	32 (31.4)	0.18 (0.08)*	41 (40.2)	0.09 (0.08)
<b>Empathy</b>						
Empathy: Fear	3 (2.9)	–	54 (52.9)	–	8 (7.8)	–
Empathy: Upset	6 (5.9)	–	34 (33.3)	–	16 (15.7)	–
Empathy: Broken Trust	1 (1.0)	–	2 (2.0)	–	61 (59.8)	–
Validation	4 (3.9)	–	43 (42.2)	–	45 (44.1)	–
Empathy: Sad	0 (0)	–	0 (0)	–	2 (2.0)	–
Empathy: Self-blame	0 (0)	–	0 (0)	–	0 (0)	–
Empathy: Time to Process	0 (0)	–	0 (0)	–	2 (2.0)	–
Grouping Total	12 (11.8)	0.05 (0.13)	79 (77.5)	0.21 (0.1)*	89 (87.3)	0.36 (0.11)**
<b>Physician Feelings</b>						
Feel: personal	4 (3.9)	–	3 (2.9)	–	7 (6.9)	–
Feel: remorse	0 (0)	–	2 (2.0)	–	1 (1.0)	–
Grouping Total	4 (3.9)	0.43 (0.2)*	5 (4.9)	0.47 (0.18)*	8 (7.8)	–0.04 (0.15)
<b>Alignment</b>						
Align with patient	5 (4.9)	–	26 (25.5)	–	50 (49.0)	–
Rebuilding partnership	2 (2.0)	–	4 (3.9)	–	45 (44.1)	–
Grouping Total	7 (6.9)	–0.29 (0.16) <sup>+</sup>	29 (28.4)	0.19 (0.09)*	73 (71.6)	0.24 (0.09)**
<b>Next steps</b>						
Next steps: General	21 (20.6)	–	40 (39.2)	–	23 (22.6)	–
Next steps: Expedite care	3 (2.9)	–	14 (13.7)	–	8 (7.8)	–
Next steps: Specific test	10 (9.8)	–	22 (21.6)	–	15 (14.7)	–
Next steps: New PCP	0 (0)	–	1 (1.0)	–	32 (31.4)	–
Next steps: Discuss	3 (2.9)	–	5 (4.9)	–	5 (4.9)	–
Next Steps: Recommendation	1 (1.0)	–	3 (2.9)	–	1 (1.0)	–
Grouping Total	32 (31.4)	0.12 (0.09)	47 (46.1)	0.18 (0.07)*	58 (56.9)	0.17 (0.08)*
<b>System improvement</b>						
Prevention: General	9 (8.8)	–	14 (13.7)	–	45 (44.1)	–
Investigation	10 (9.8)	–	17 (16.7)	–	21 (20.6)	–
Error reporting	0 (0)	–	0 (0)	–	4 (3.9)	–
Prevention: specific	0 (0)	–	0 (0)	–	6 (5.9)	–
Grouping Total	14 (13.7)	0.34 (0.11)**	23 (22.6)	0.14 (0.09)	50 (49.0)	0.13 (0.08)
<b>Financial</b>						
Financial concerns: Acknowledge	0 (0)	–	0 (0)	–	0 (0)	–
Financial concerns: Address	0 (0)	–	0 (0)	–	1 (1.0)	–
Grouping Total	0 (0)	–	0 (0)	–	0 (0)	–0.38 (0.37)

(Continued)

TABLE 2 Continued

Group Code	Vignette 1		Vignette 2		Vignette 3	
	N (%) Present	$\beta$ (SE)	N (%) Present	$\beta$ (SE)	N (%) Present	$\beta$ (SE)
<b>General communication skills</b>						
Use patient's name	44 (43.1)	0.1 (0.08)	26 (25.5)	0.03 (0.09)	25 (24.5)	0 (0.09)
<b>General communication skills</b>						
Use Jargon	46 (45.1)	0.02 (0.08)	6 (5.9)	−0.26 (0.18)	2 (2.0)	0.01 (0.28)
<b>Rationalization</b>						
Rationalize error	0 (0)	–	0 (0)	–	6 (5.9)	−0.19 (0.17)
<b>Minimization</b>						
Minimization	0 (0)	–	5 (4.9)	−0.24 (0.17)	0 (0)	–
<b>Address mammogram results</b>						
Prior results	78 (76.5)	−0.03 (0.1)	6 (5.9)	0 (0.16)	2 (2.0)	−0.35 (0.26)
Connect to current concern	58 (56.9)	0.01 (0.08) <sup>+</sup>	8 (7.8)	0.13 (0.15)	0 (0)	–

$\beta$  indicates the strength and direction of the relationship between for presence of each code and layperson rating response as found by the multiple linear regression analysis.

<sup>+</sup>Indicates a  $p$  value between 1 and .05.

<sup>\*</sup>Indicates significance at  $p < .05$ .

<sup>\*\*</sup>Indicates significance at  $p < .01$ .

<sup>\*\*\*</sup>Indicates significance at  $p < .001$ .

## Acknowledgement that information was missed by the care team

Approximately three quarters of residents ( $N = 76$ , 74.5%) generally acknowledged that information was missed by the care team or provided some explanation as to why information was missed in at least one of their responses. For example, one resident said, “I’m not sure how that was missed”, where another said, “I think I must have overlooked it when I was reviewing your chart prior to our last visit.”

## Acknowledgement of a diagnostic delay

More than half of the residents ( $N = 59$ , 57.8%) acknowledged a diagnostic delay or a delay in care in at least one of the three vignettes. For example, one resident stated, “We could have known about this about a year ago,” while another remarked, “... your mammogram did indicate last year that there was a calcification in the area, and we should have ordered the biopsy for that.” Very few responses ( $N = 5$ , 4.9%) minimized the clinical impact of the time delay. For example, one resident said, “It may not necessarily be that this year would’ve made a huge difference.”

## An apology

Nearly all ( $N = 92$ , 90.2%) of residents provided responses which included either a general apology or a specific apology in at least one of the three vignettes. For example, one resident said “I’m very sorry about this” and another said “I’m very sorry that I didn’t tell you about this last year.”

## Error disclosure and personal responsibility

More than half ( $N = 66$ , 64.7%) of residents provided responses which included some admission that an error did occur, to at least one of the three Vignettes. For example, one resident said, “I apologize for the mistake that happened”, another said “It’s absolutely my fault for not following up on that and sending you

for a biopsy” and another said “I’d like to start by saying that a mistake was made on your mammogram last year.” Rationalization was uncommon, but a minority ( $N = 6$ , 5.9%) of resident responses to vignette 3 attempted to normalize the mistake. One resident said, “although we are physicians and we do go through extensive training, we do make a mistake from time to time and after all, we are we all human” and another said, “As a human, it is unfortunately hard to escape making mistakes.”

## Empathy

Nearly all ( $N = 98$ , 96.1%) residents provided at least one response expressing a component of empathy. Manifestations of empathy included validating or acknowledging the patient’s emotions, acknowledging that trust was broken, and offering time to process the information. For example, one resident acknowledged the patient’s fears by saying “I recognize that this can be really scary”. Another resident acknowledged that trust was broken by saying “it’s completely understandable if you feel like you can’t trust us anymore.”

## Personal feelings or remorse

Few ( $N = 14$ , 13.7%) residents provided responses which mentioned their own personal feelings to at least one of the three Vignettes. For example, one resident expressed how they were affected by saying “I’m also very upset about this” and another expressed remorse by saying “I regret that this happened.”

## Patient alignment

Many ( $N = 77$ , 75.5%) residents provided at least one response expressing alignment with the patient through commitments of support, working together as a team, or an intention to regaining trust. One resident demonstrated alignment with the patient by saying “I would like to do everything I can to help you through this, and I would like to try to support you moving forward.”

Another addressed trust by saying “*I think my hope is that over time, if we continue to work together, that I can slowly rebuild your trust in me.*” Only one resident addressed the financial aspect of the error by saying “*we can do whatever we can at this point in time financially to help resolve our error for you.*”

### Patient-centered communication

Approximately half ( $N=56$ , 54.9%) of residents used the patient’s name in at least one of their responses. Regarding use of medical jargon, about half ( $N=47$ , 46.1%) of residents used the term “*calcification*” without further explanation about what it meant or its clinical significance. This behavior was mostly present in responses to Vignette 1.

### Plans for next steps in care

Many ( $N=79$ , 77.5%) residents provided responses which contained mentions of next steps in care, to at least one of the three Vignettes. Responses ranged from general next steps to specific next steps, such as recommending a biopsy. Some responses made explicit recommendations for next steps or expressed that they would be expedited. Some also mentioned that they would welcome further discussion regarding the error or expressed understanding of the desire to transition to a new provider. For example, one resident said “*let’s repeat your mammogram imaging now and let’s get a biopsy of that area, as well*” and another said “*I will do everything in my power to get you in for biopsies as soon as possible to make sure that this is taken care of.*” One resident mentioned the possibility of switching to a new provider by saying “*Alternatively, if you don’t feel comfortable following with my group anymore, I can certainly refer you to some well-respected primary care physicians in the area to establish care.*”

### Plans to prevent future mistakes

Many ( $N=62$ , 67.8%) residents provided responses which made mention of preventing future mistakes from happening again to at least one of the three Vignettes. Responses ranged from general desires to prevent mistakes to specific and concrete action items. They also mentioned plans to investigate what happened and to report to hospital systems. For example, one resident vaguely mentioned “*I will ensure that this doesn’t happen again*” whereas another was more specific by saying “*Going forward, I will be sure to review every test result that I have with you personally.*” Another resident said, “*we’re going to start an investigation to see what had happened.*”

## Thematic elements predictive of layperson ratings

The results of the three multiple linear regression analyses are presented in Table 2. For Vignette 1, communication themes predictive of layperson ratings on responses included Apology ( $\beta=.34$ ,  $p<.001$ ), accepting responsibility ( $\beta=.48$ ,  $p=.001$ ), physician’s feelings ( $\beta=.43$ ,  $p=.032$ ), and system improvement ( $\beta=.34$ ,  $p=.003$ ) were significant predictors of layperson ratings.

These elements accounted for 49% of the variance (adjusted  $R^2=.49$ ),  $F(13, 88)=8.49$ ,  $p<.001$ .

For Vignette 2, communication themes predictive of layperson ratings on responses were acknowledging delays ( $\beta=.19$ ,  $p<.014$ ), accepting responsibility ( $\beta=.18$ ,  $p=.038$ ), empathy ( $\beta=.21$ ,  $p=.032$ ), physician feelings ( $\beta=.47$ ,  $p=.012$ ), alignment ( $\beta=.19$ ,  $p=.028$ ), and next steps ( $\beta=.18$ ,  $p=.019$ ) were significant predictors of layperson ratings. These elements accounted for 32% of the variance (adjusted  $R^2=.32$ ),  $F(14, 87)=4.35$ ,  $p<.0001$ .

For Vignette 3, communication themes predictive of layperson ratings included empathy ( $\beta=.36$ ,  $p=.002$ ), alignment ( $\beta=.24$ ,  $p=.006$ ), and next steps ( $\beta=.17$ ,  $p=.028$ ) were significant predictors of layperson ratings. These elements accounted for 28% of the variance (adjusted  $R^2=.16$ ),  $F(14, 87)=2.42$ ,  $p=.007$ .

## Concordance between expert recommendations and thematic groups predictive of layperson ratings

Regarding concordance between experts and laypeople, we found significant associations between layperson ratings and 7 of the 10 expert-recommended thematic groups. Of the expert recommended thematic groups significant predictors of layperson ratings were acknowledge delays in vignette 2 ( $\beta=.19$ ,  $p<.014$ ), apology in vignette 1 ( $\beta=.34$ ,  $p<.001$ ), accept responsibility in vignette 1 ( $\beta=.48$ ,  $p=.001$ ) and 2 ( $\beta=.18$ ,  $p=.038$ ), empathy in vignette 2 ( $\beta=.21$ ,  $p=.032$ ) and 3 ( $\beta=.36$ ,  $p=.002$ ), alignment in vignette 3 ( $\beta=.24$ ,  $p=.006$ ), next steps in vignette 2 ( $\beta=.18$ ,  $p=.019$ ) and 3 ( $\beta=.17$ ,  $p=.028$ ), and system improvements in vignette 1 ( $\beta=.34$ ,  $p=.003$ ). Though experts recommended acknowledging missed results, addressing mammogram results and addressing finances, they were not significantly predictive of layperson ratings in any vignette. Of the three thematic groups experts recommended avoiding in error disclosure (minimization, rationalization and jargon), only jargon appeared with regular frequency (45% of responses in vignette 1). However, the presence of jargon was not negatively associated with layperson ratings.

## Discussion

Our analysis of spoken responses to a hypothetical patient with a delayed diagnosis of breast cancer demonstrates significant variability in how resident physicians would approach this communication challenge. While essentially all respondents would openly address the fact that the mammogram was abnormal a year ago, residents did not reliably provide other key information, expressions of accountability, or follow-up commitments expected by patients. For example, across the whole case, only 57.8% of residents openly acknowledged that the care was delayed, and 67.8% expressed a plan to prevent future errors. These findings highlight the need to develop curricular material and practice tools that prepare each individual learner for comprehensive, open and effective error disclosure



conversations. A related implication is that assessment and feedback systems should provide feedback to reinforce key behaviors that were displayed, and reminders to practice behaviors that were needed but omitted.

Our linear regression analysis provided useful insights into the behaviors that most strongly influence how laypeople rate a physician's error disclosure skills. The strongest associations were expressions of accountability (0.48), personal regret (0.47), apology (0.34), and intentions to prevent future mistakes (0.34). These behaviors varied across vignettes, corresponding to the information or response appropriate to the patient's prompt. This reflects that effective error disclosure is both an art and a science, in which the physician must both cover certain key topics, while connecting them in a clear and sensible way. The importance of these top 4 behaviors largely aligns with prior literature emphasizing the importance of accountability and apology, although the positive contribution of the physician's personal feelings is less expected (32). This could indicate that laypeople felt as though residents aligned with them through shared expression of being personally upset about the case. Although it was a relatively rare behavior (13.7% of residents), more investigation is warranted to understand how physicians can effectively express their personal feelings without taking attention away from the patient's concerns.

Several relatively uncommon communication behaviors may warrant emphasis by medical educators using the VCA or other simulations. First, only one respondent spontaneously addressed the potential financial implications of the medical error, which is a concern for many patients after medical harm. While it was likely appropriate that residents deprioritized this topic in a simulation of the early stages of disclosure, its absence suggests educators should also use simulation cases designed to elicit raising or responding to this topic as it is appropriate for providers to be prepared to respond if a patient brings up financial implications. Second, a small percentage of residents used rationalization or minimization behaviors that would likely anger or alienate patients. Responses with these behaviors were associated with negative beta-coefficients, although this finding did not reach statistical significance. Remediating this uncommon, but critical, behavior would require other approaches, such as a faculty coach (33) or using artificial intelligence to create individualized feedback (34).

Although nearly all residents offered either a general or specific apology at some point, they were spread across vignettes (e.g., 58.9% of vignette 1 responses included an apology, and 69.6% in vignette 2). Yet, apologies only appeared to influence the overall rating most when done at the outset of the conversation, suggesting that even brief delays can reduce the perceived quality of the apology. This finding warrants further investigation. Although prior literature has focused on the overall presence of an apology or its phrasing (14, 35), it may be that the timing is more relevant than previously appreciated.

The linear regression analysis identified transcript-based variables that explained up to 49% of the variance in communication performance. This is robust but suggests that factors other than word choice also contribute to how laypeople

rate the error disclosure skills of physicians. This may include prosodic features such as intonation, volume, pace, and stress. Although there are ways to measure these features of speech, they fall outside the usual expertise of medical educators and may be harder to coach than language and thematic content. Yet, this suggests an opportunity to utilize both speech and language metrics to fully understand error disclosure effectiveness.

A secondary aim of this study was to examine the alignment between error disclosure content recommended by experts and the communication behaviors that contribute to higher layperson ratings of disclosure. We found that all the expert-recommended thematic groups were associated with a positive beta coefficient, and all except three (acknowledge missed results, addressing mammogram results and addressing finances) were significantly associated with higher layperson ratings. This adds to the validity evidence for the VCA as an assessment tool, including prior work showing that laypeople are reliable surrogates for patients who were injured by errors (29). Contrary to expert expectations, jargon was not negatively associated with layperson ratings. Despite this result, existing literature shows that jargon can hinder patient comprehension and shared decision-making (36, 37). These findings highlight the importance of integrating both patient perspectives and expert input when identifying optimal physician communication strategies.

Our study has important limitations. First, we analyzed responses to a simulated scenario, which may elicit different behaviors from real-life disclosure. For example, our results may overestimate the portion of residents who apologize or offer open explanations of the error, compared with high-stakes real scenarios. Second, the VCA does not evaluate body language or facial expressions that contribute to communication efficacy in real life, and our coding was based on transcripts alone, which excludes the contributions from tone or speaking style. Third, the VCA also does not measure encourage open-ended questions or silence, although curiosity and patience are desirable hallmarks of empathy that were not assessed in this study. Fourth, our codebook may not have captured all relevant behaviors that can be identified from text. Fifth, we did not ask layperson raters why they provided specific ratings; assessing their emotional responses and viewpoints could offer novel insight. Sixth, the sample represents residents who participated in research studies at a mix of internal medicine and family medicine residencies; findings may not generalize to other specialties and non-participants may have different communication behaviors. Lastly, this study analyzed physician-patient communication alone, and did not investigate the important role of other healthcare professions in team-based disclosure.

## Conclusion

Resident physicians vary in which components they include in error disclosure, missing opportunities to meet patient expectations. Uncommonly, some residents employed rationalization or minimization in their responses. Tailoring an

assessment and feedback system to encourage responders to include the themes that layperson raters value most and to omit harmful expressions represents an important feature for future software for error disclosure communication training. Expert-recommended communication behaviors also aligned highly with the themes associated with higher ratings from laypeople, adding evidence for the validity of using VCA for error disclosure skill assessment.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

## Ethics statement

The studies involving humans were approved by The University of Washington Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because all users signed a software license consent that governed data use. They did not provide written consent for study participation, but did provide verbal consent (and could opt out at any time by stopping participation).

## Author contributions

EG: Writing – original draft, Writing – review & editing. AD: Writing – original draft, Writing – review & editing. AK: Writing – review & editing. TG: Writing – review & editing. KM: Writing – review & editing. AW: Writing – original draft, Writing – review & editing.

## References

1. Mazor KM, Simon SR, Gurwitz JH. Communicating with patients about medical errors: a review of the literature. *Arch Int Med.* (2004) 164(15):1690–7. doi: 10.1001/archinte.164.15.1690
2. National Quality Forum. *Safe Practices for Better Healthcare*. Washington, DC, USA: National Quality Forum (2006).
3. Gallagher TH, Kachalia A. Responding to medical errors—implementing the modern ethical paradigm. *N Engl J Med.* (2024) 390(3):193–7. doi: 10.1056/NEJMp2309554
4. Gallagher TH, Garbutt JM, Waterman AD, Flum DR, Larson EB, Waterman BM, et al. Choosing your words carefully: how physicians would disclose harmful errors to patients. *Arch Int Med.* (2006) 166(15):1585–93. doi: 10.1001/archinte.166.15.1585
5. Delbanco T, Bell SK. Guilty, afraid, and alone—struggling with medical error. *NEJM.* (2007) 357(17):1682–3. doi: 10.1056/NEJMp078104
6. Gallagher TH, Waterman AD, Garbutt JM, Kapp JM, Chan DK, Dunagan CW, et al. US And Canadian physicians' attitudes and experiences regarding disclosing

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Several authors of this work received grant funding from the NBME.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frhs.2025.1577092/full#supplementary-material>

errors to patients. *Arch Int Med.* (2006) 166(15):1605–11. doi: 10.1001/archinte.166.15.1605

7. White AA, Bell SK, Krauss MJ, Garbutt MJ, Dunagan CW, Fraser VJ, et al. How trainees would disclose medical errors: educational implications for training programs. *Med Educ.* (2011) 45(4):372–80. doi: 10.1111/j.1365-2923.2010.03875.x

8. White AA, Gallagher TH, Krauss MJ, Garbutt J, Waterman AD, Dunagan CW, et al. The attitudes and experiences of trainees regarding disclosing medical errors to patients. *Acad Med.* (2008) 83(3):250–6. doi: 10.1097/ACM.0b013e3181636e96

9. Accreditation Council for Graduate Medical Education. ACGME Common Program Requirements (Residency): section IV.B.1.f.(1).(g). Available at: [https://www.acgme.org/globalassets/pfassets/programrequirements/cprresidency\\_2023.pdf](https://www.acgme.org/globalassets/pfassets/programrequirements/cprresidency_2023.pdf) (Accessed January 2, 2025).

10. Centers for Medicare & Medicaid Services. "The Patient Safety Structural Measure's Five Domain Attestations" in FY 2025 IPPS/LTCH PPS, 89 FR 69486, Vol 42. Available at: <https://www.federalregister.gov/d/2024-17021/page-69486> (Accessed January 3, 2025).

11. Iedema R, Sorensen R, Manias E. Patients' and family members' experiences of open disclosure following adverse events. *Int J Qual Health Care*. (2008) 20(6):421–32. doi: 10.1093/intqhc/mzn043
12. Shapiro J, Robins L, Galowitz P, Gallagher TH, Bell S. Disclosure coaching: an ask-tell-ask model to support clinicians in disclosure conversations. *J Patient Saf*. (2021) 17(8):e1364–70. doi: 10.1097/PTS.0000000000000491
13. White AA, Brock DM, McCotter PI, Shannon SE, Gallagher TH. Implementing an error disclosure coaching model: a multicenter case study. *J Healthc Risk Manag*. (2017) 36(3):34–45. doi: 10.1002/jhrm.21260
14. Prentice JC, Bell SK, Thomas EJ, Schneider EC, Weingart SN, Weissman JS, et al. Association of open communication and the emotional and behavioural impact of medical error on patients and families: state-wide cross-sectional survey. *BMJ Qual Saf*. (2020) 29(11):883–94. doi: 10.1136/bmjqs-2019-010367
15. Myren BJ, de Hullu JA, Bastiaans S, Kokks JJ, Hermens RPMG, Zusterzeel PLM. Disclosing adverse events in clinical practice: the delicate act of being open. *Health Commun*. (2022) 37(2):191–201. doi: 10.1080/10410236.2020.1830550
16. Loren DJ, Klein EJ, Garbutt J, Krauss MJ, Fraser V, Dunagan WC, et al. Medical error disclosure among pediatricians: choosing carefully what we might say to parents. *Arch Pediatr Adolesc Med*. (2008) 162(10):922–7. doi: 10.1001/archpedi.162.10.922
17. Dhawale T, Zech J, Greene SM, Roblin DW, Brigham KB, Gallagher TH, et al. We need to talk: provider conversations with peers and patients about a medical error. *J Patient Saf Risk Manag*. (2019) 24(4):140–6. doi: 10.1177/2516043519863578
18. Reisch LM, Prouty CD, Elmore JG, Gallagher TH. Communicating with patients about diagnostic errors in breast cancer care: providers' attitudes, experiences, and advice. *Patient Educ Couns*. (2020) 103(4):833–8. doi: 10.1016/j.pec.2019.11.022
19. Crimmins AC, Wong AH, Bonz JW, Tsyrlunik A, Jubanyik K, Dziura JD, et al. "To err is human" but disclosure must be taught: a simulation-based assessment study. *Simul Healthc*. (2018) 13(2):107–16. doi: 10.1097/SIH.0000000000000273
20. Stroud L, McIlroy J, Levinson W. Skills of internal medicine residents in disclosing medical errors: a study using standardized patients. *Acad Med*. (2009) 84(12):1803–8. doi: 10.1097/ACM.0b013e3181bf9fef
21. Hannawa AF. Shedding light on the dark side of doctor-patient interactions: verbal and nonverbal messages physicians communicate during error disclosures. *Patient Educ Couns*. (2011) 84(3):344–51. doi: 10.1016/j.pec.2011.04.030
22. Chan DK, Gallagher TH, Reznick R, Levinson W. How surgeons disclose medical errors to patients: a study using standardized patients. *Surgery*. (2005) 138:851–8. doi: 10.1016/j.surg.2005.04.015
23. Mazor KM, King AM, Hoppe RB, Kochersberger AO, Yan J, Reim JD. Video-Based communication assessment: development of an innovative system for assessing clinician-patient communication. *JMIR Med Educ*. (2019) 5(1):e10400. doi: 10.2196/10400
24. Mazor KM, King AM, Hoppe RB, D'Addario A, Musselman TG, Tallia AF, et al. Using crowdsourced analog patients to provide feedback on physician communication skills. *Patient Educ Couns*. (2021) 104(9):2297–303. doi: 10.1016/j.pec.2021.02.047
25. White AA, King AM, D'Addario AE, Brigham KB, Dintzis S, Fay EE, et al. Effects of practicing with and obtaining crowdsourced feedback from the video-based communication assessment app on resident Physicians' adverse event communication skills: pre-post trial. *JMIR Med Educ*. (2022) 8(4):e40758. doi: 10.2196/40758
26. White AA, King AM, D'Addario AE, Brigham KB, Bradley JM, Gallagher TH, et al. Crowdsourced feedback to improve resident physician error disclosure skills: a randomized clinical trial. *JAMA Netw Open*. (2024) 7(8):e2425923. doi: 10.1001/jamanetworkopen.2024.25923
27. Kolb DA. *Experiential Learning: Experience as the Source of Learning and Development*. NJ: Prentice-Hall (1984).
28. Faro JM, D'Addario A, King AM, Mazor KM, Pbert L, Sadasivam RS, et al. Video-based communication assessment for weight management counseling training in medical residents: a mixed methods study. *BMC Med Educ*. (2022) 22(1):899. doi: 10.1186/s12909-022-03984-6
29. White AA, King AM, D'Addario AE, Brigham KB, Dintzis S, Fay EE, et al. Video-based communication assessment of physician error disclosure skills by crowdsourced laypeople and patient advocates who experienced medical harm: reliability assessment with generalizability theory. *JMIR Med Educ*. (2022) 8(2):e30988. doi: 10.2196/30988
30. Vriesema CC, Gehlbach H. Assessing survey satisficing: the impact of unmotivated questionnaire responding on data quality. *Educ Res*. (2021) 50(9):618–27. doi: 10.3102/0013189X211040054
31. Agency for Healthcare Research and Quality. "AHRQ Communication and Optimal Resolution Toolkit (CANDOR). Available at: <https://www.ahrq.gov/patient-safety/settings/hospital/candor/modules/guide5/notes.html> (Accessed April 30, 2025).
32. Mazor KM, Greene SM, Roblin D, Lemay CA, Firreno CL, Calvi J, et al. More than words: patients' views on apology and disclosure when things go wrong in cancer care. *Patient Educ Couns*. (2013) 90(3):341–6. doi: 10.1016/j.pec.2011.07.010
33. Grossniklaus E, King AM, D'Addario AE, Brigham KB, Gallagher TH, Musselman TG, et al. Incorporating the Video Communication Assessment for Error Disclosure in Residency Curricula: A Mixed Methods Study of Faculty Perceptions. Submitted to *Frontiers in Health Services Research, Patient Safety Section*.
34. Lin IW, Sharma A, Rytting C, Miner AS, Suh J, Althoff T. IMBUE: improving interpersonal effectiveness through simulation and just-in-time feedback with human-language model interaction. *Annual Meeting of the Association for Computational Linguistics* (2024).
35. Redelmeier DA, Roach J. Psychology insights on apologizing to patients. *J Hosp Med*. (2024):1–4. doi: 10.1002/jhm.13585
36. Castro CM, Wilson C, Wang F, Schillinger D. Babel babble: physicians' use of unclarified medical jargon with patients. *Am J Health Behav*. (2007) 31(1):S85–95. doi: 10.5993/AJHB.31.s1.11
37. Links AR, Callon W, Wasserman C, Walsh J, Beach MC, Boss EF. Surgeon use of medical jargon with parents in the outpatient setting. *Patient Educ Couns*. (2019) 102(6):1111–8. doi: 10.1016/j.pec.2019.02.002